PennState
Eberly College of Science

Ashtekar Frontiers of Science

The 2026 Ashtekar Frontiers of Science lecture series will explore how researchers in the Eberly College of Science approach some of today's biggest societal questions using different approaches. The 2026 lecture series brings together researchers from disparate disciplines to showcase how faculty from across the college's seven departments leverage wide-ranging methodologies and theoretical approaches to tackle pressing scientific issues and help shape the future of science, health, and humanity.

Scan the QR code to subscribe to our mailing list.

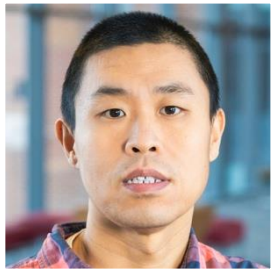
The lecture series was founded by Abhay Ashtekar, founding director of the Institute for Gravitation and the Cosmos and a member of the National Academy of Sciences. It owes its success to Barbara Kennedy, who presided over the series during its first 25 years, making it one of the most successful science outreach events in central Pennsylvania.

001 Chemical and Biomedical Engineering Building

"Origins"
January 24, 2026


Suvrath Mahadevan | Tom Stewart

"Medicine"
January 31, 2026

Santhosh Girirajan | Wenrui Hao

"Climate"
February 7, 2026


Helen Greatrex | John Harlim

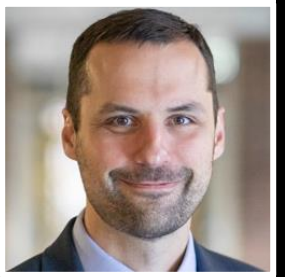
"Materials"
February 14, 2026


Bryce Gadway | Danielle Hickey

"Resources"
February 21, 2026


Charlie Anderson | Joey Cotruvo

"Education"
February 28, 2026

Jennelle Malcos | Matthew Beckman

Unexpected Pairings Addressing Big Questions

PennState Eberly College of Science | Ashtekar Frontiers of Science

AI Support for Teaching and Learning (at scale)

Matthew D. Beckman
Department of Statistics

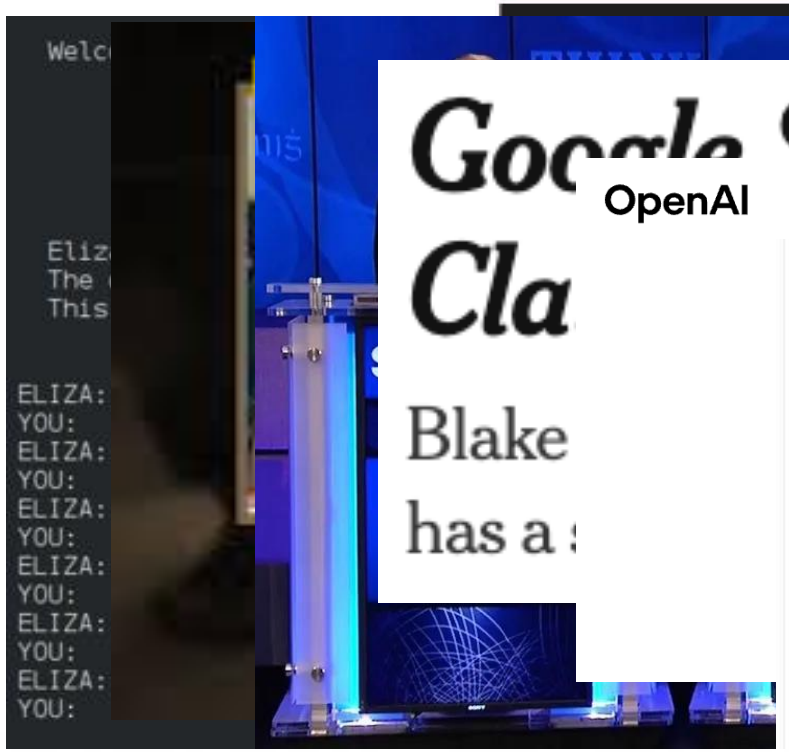
February 28, 2026

Resource Page

Unexpected Pairings Addressing Big Questions

PennState Eberly College of Science | Ashtekar Frontiers of Science

The learning environment *matters*



1966 1997 2011 2022 2023 2026 Today

Unexpected Pairings Addressing Big Questions

PennState Eberly College of Science | Ashtekar Frontiers of Science

JOURNAL OF STATISTICS AND DATA SCIENCE EDUCATION
2026, VOL. 34, NO. 1, 4–15
https://doi.org/10.1080/26939169.2025.2596684

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Students' Statistical Thinking When Using Generative AI: A Descriptive Case Study

V. N. Vimal Rao, Amos Jeng, Julianna Drew, Bhuvan Kala, and Sanjana Gongati

University of Illinois at Urbana Champaign, Urbana, IL

How AI Impacts Skill Formation

Judy Hanwen Shen* Alex Tamkin†

ANTHROP\C

February 3, 2026

Abstract

AI assistance produces significant productivity gains across professional domains, particularly for novice workers. Yet how this assistance affects the development of skills required to effectively supervise AI remains unclear. Novice workers who rely heavily on AI to complete unfamiliar tasks may compromise their own skill acquisition in the process. We conduct randomized experiments to study how developers gained mastery of a new asynchronous programming library with and without the assistance of AI. We find that AI use impairs conceptual understanding, code reading, and debugging abilities, without delivering significant efficiency gains on average. Participants who fully delegated coding tasks showed some productivity improvements, but at the cost of learning the library. We identify six distinct AI interaction patterns, three of which involve cognitive engagement and preserve learning outcomes even AI assistance. Our findings suggest that AI-enhanced productivity is not a shortcut to competence and AI assistance should be adopted into workflows to preserve skill formation – particularly in safety-critical domai

Where's Matt?

Never with AI ⟷ Everything with AI

Lessons for Education

- Not an "end of homework"
- Teaching shifts emphasis
- Access and Scale Improve
- Caution: "ELIZA Effect"

Personalized Practice / Tutoring

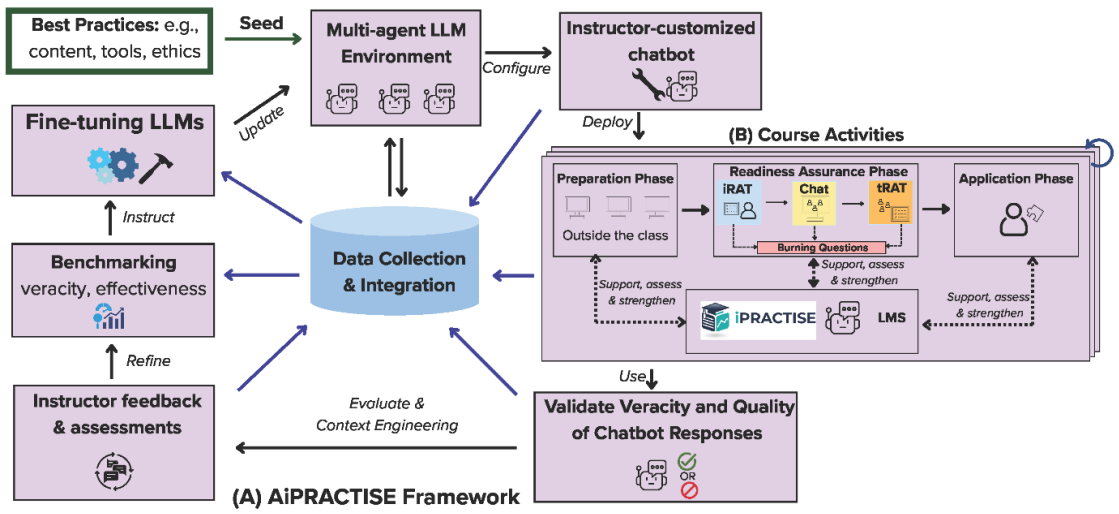


Fig. Schematic representation of Ai-Practise Project related technologies.

- Ai-Practise
  - SSRI AI Seed Grant; Sy-Miin Chow, PI
  - Extends iPractise & T-Practise NSF Grants
- SPARK Self-Regulated Learning
  - Rayne Sperling, PI
- Context-aware Chatbots for Education
  - Khan Academy's Khanmigo
  - Pearson AI Study Tool
  - NSF-Funded HaLLMos Tool (https://hallmos.com)

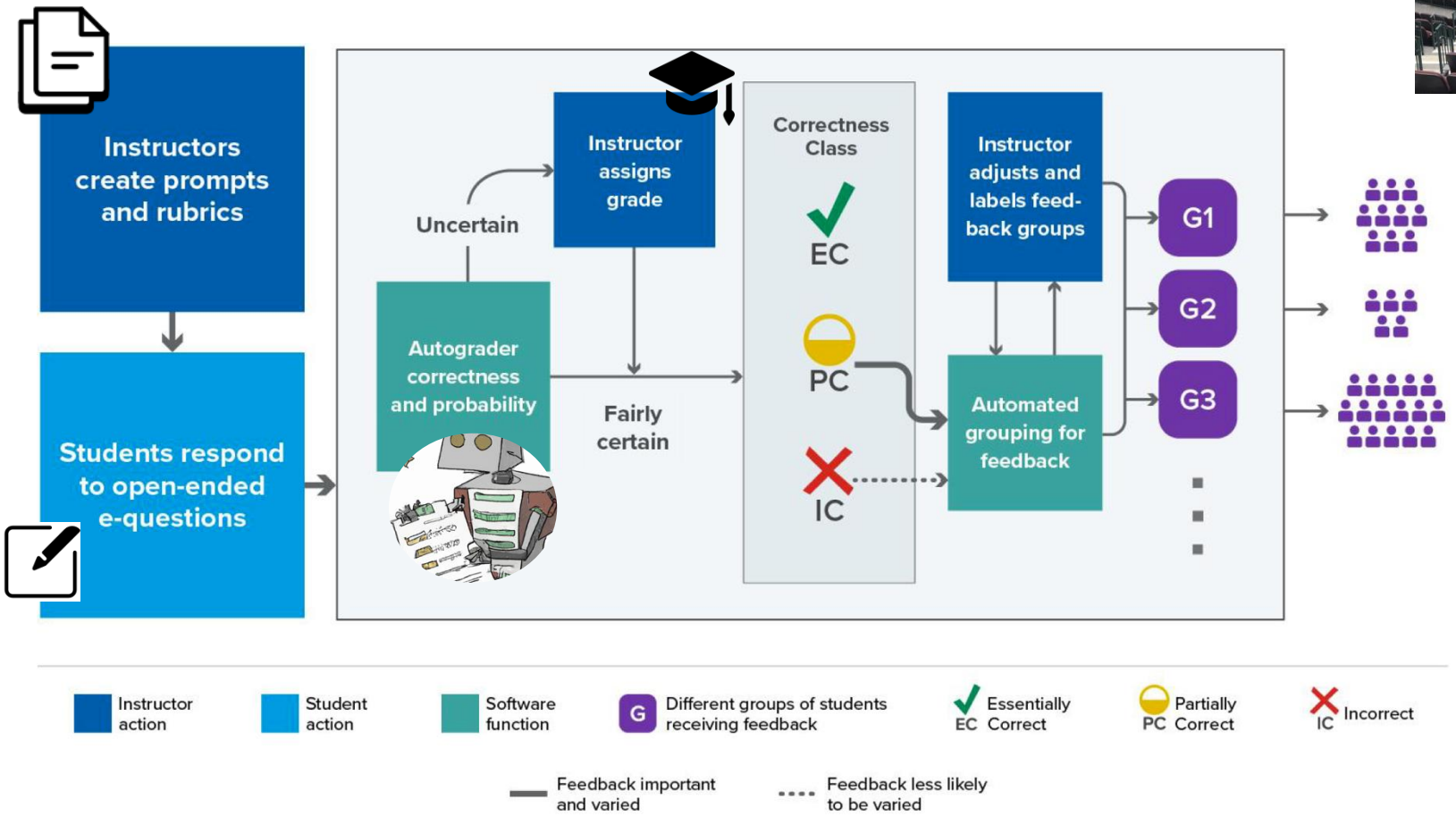
AI-supported "Check-for-Understanding"

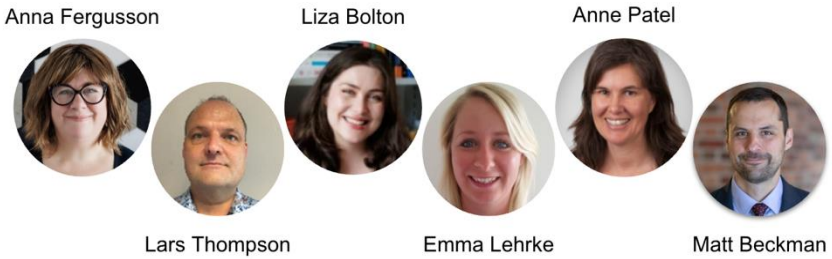
Matt Beckman
(Statistics)

Becky Passonneau
(Computer Science)

Dennis Pearl
(Statistics)

PennState





Waipapa Taumata Rau University of Auckland
+
PennState

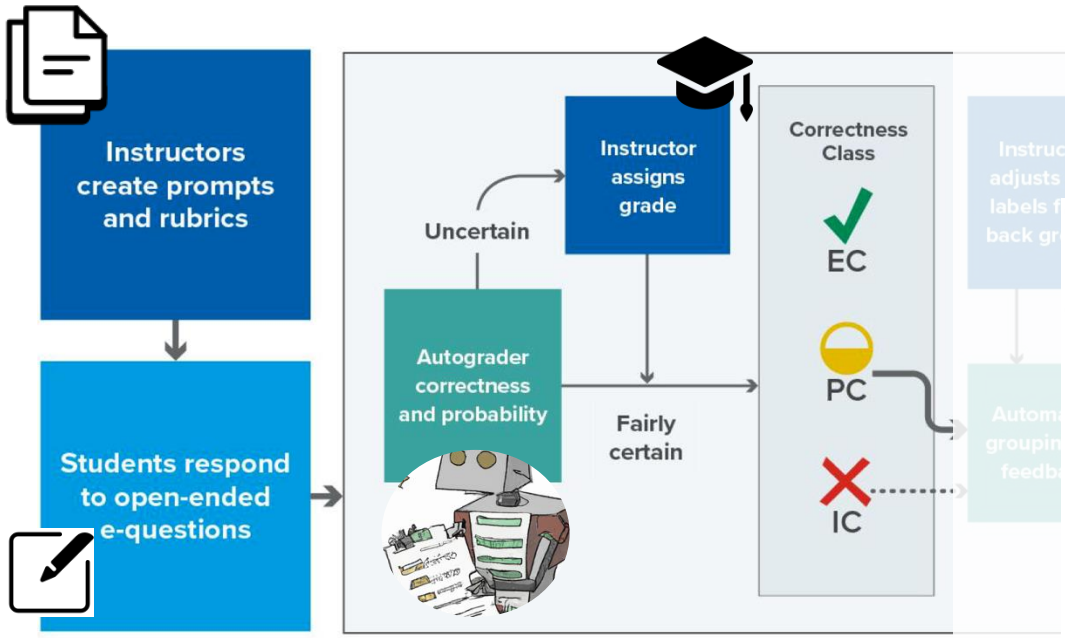
Prototype:
Dr. Anna Fergusson
University of Auckland (NZ)

Quick, Write!

Question: How would you describe the value of this kind of formative assessment? For students? For Instructors?

(Seriously, just share whatever comes to mind; we'll use them in a moment)

Does it… *work?*



"Scoring" Agreement:

- Instructors $\approx 0.7 - 0.8$
- TA's $\approx 0.6 - 0.7$
- Our custom AI ≈ 0.7

Human-in-the-Loop Combination:

- $0.85 - 0.9$ *with* deferral control

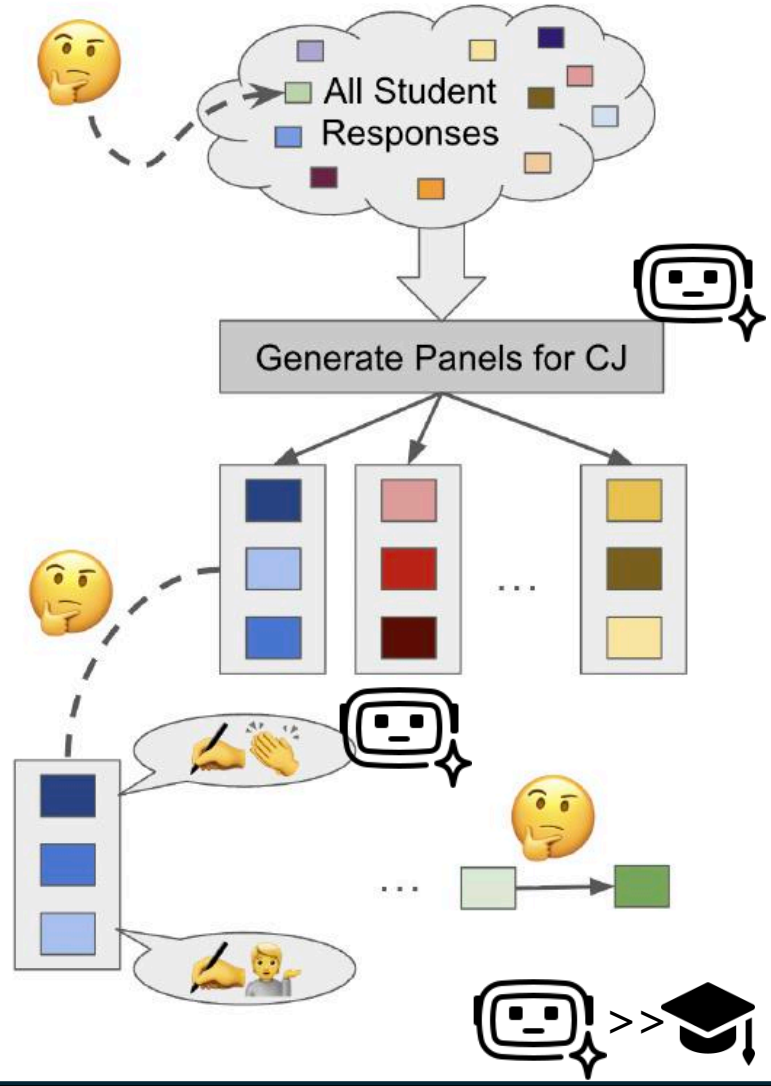
Threshold	Deferral Rate	Simulated HIL Accuracy
0.68	9.5%	0.855
0.75	13.2%	0.861
0.80	16.0%	0.871
0.85	20.2%	0.884
0.90	25.6%	0.899

Students: "*What's in it for me??*"

(translation: I thought he said *formative* assessment)

- Comparative Judgments (CJ)
- Evaluate other responses
- Give (& get) feedback

p.s. it's the same QR code as "Quick, Write!" *so you can either scan OR just refresh the page.*



Prototype: Dr. Anna Fergusson (University of Auckland, NZ)


Matt Beckman


Wenpeng Yin


ChanMin Kim


Neil Hatfield


Anna Fergusson

PennState + Waipapa Taumata Rau University of Auckland

And then what?

- How do LLM / AI models **handle linguistic diversity**?
- **New learning theory** for formative comparative judgement (CJ)?
- How does AI-supported CJ influence **critical thinking**?
- **Clustering** LLM representations for student feedback?
- **Constraints** on task & rubric design/implementation?

Acknowledgments

- US National Science Foundation (NSF DUE-2236150; DUE-2417294)
- Penn State Center for Socially Responsible Artificial Intelligence (CSRAI)
- Penn State Social Science Research Institute (SSRI) AI Seed Grant
- Strategic partnership between University of Auckland and Penn State University
- Students, instructors, and institutional partners that contributed data and permitted us to use it for our research (Penn State IRB: Study 00022615)

References

- Anthropic. (2026, January 22). Claude's new constitution. https://www.anthropic.com/news/claude-new-constitution
- Beckman, Burke, Fiochetta, Fry, Lloyd, Patterson, Tang (2024). Developing Consistency Among Undergraduate Graders Scoring Open-Ended Statistics Tasks. Preprint URL: https://arxiv.org/abs/2410.18062
- Brooks, D. (2022, February 2). Opinion: In the age of A.I., major in being human. The New York Times. https://www.nytimes.com/2023/02/02/opinion/ai-human-education.html
- Grant, N., & Metz, C. (2022, June 12). Google engineer claims chatbot AI is sentient. The New York Times. https://www.nytimes.com/2022/06/12/technology/google-chatbot-ai-blake-lemoine.html
- Li, Lloyd, Beckman, & Passonneau (2023). Answer-state Recurrent Relational Network (AsRRN) for Constructed Response Assessment and Feedback Grouping. Findings of the Association for Computational Linguistics: EMNLP 2023. https://doi.org/10.18653/v1/2023.findings-emnlp.254
- Lloyd, Beckman, Pearl, Passonneau, Li, & Wang (2022). Foundations for AI-Assisted Formative Assessment Feedback for Short-Answer Tasks in Large-Enrollment Classes. In Proceedings of the eleventh international conference on teaching statistics. Rosario, Argentina.
- Shen, J.H. & Tamkin, A. (2026). How AI Impacts Skill Formation. Preprint URL: https://doi.org/10.48550/arXiv.2601.20245
- U.S. Department of Education, Office of Educational Technology. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations. U.S. Department of Education. https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf
- Wei, Beckman, Pearl, & Passonneau (in review). Concept-based Rubrics Improve LLM Formative Assessment and Data Synthesis. https://arxiv.org/pdf/2504.03877

Additional Images

- [Ai-Practise Schematic] Sy-Miin Chow
- [Autograder Schematic] Becky Passonneau, Matthew Beckman, Dennis Pearl
- [Chat Bot Logos] https://emag.directindustry.com/2025/08/20/best-ai-chatbots-data-insights-study/
- [ChatGPT] https://openai.com/index/chatgpt/
- [Chess] https://www.facebook.com/WIONews/posts/todayinhistory-1997-ibms-deep-blue-defeats-garry-kasparov-in-chess-match/2358635227680432/
- [Classroom Photos] Matthew Beckman
- [Comparative Judgment Schematic] Matthew Beckman, Anna Fergusson
- [ELIZA] https://en.wikipedia.org/wiki/ELIZA
- [Elon Musk Tweet] https://www.bbc.co.uk/newsround/64210050.amp
- [Google Photos Screenshot (& Adaptation)] Matthew Beckman
- [Grading Robot] Dall-E
- [History of eLearning] https://elearningindustry.com/history-of-elearning-infographic-education-2012
- [Jeopardy] https://blog.ted.com/how-did-supercomputer-watson-beat-jeopardy-champion-ken-jennings-experts-discuss/
- [Icons for "Edit" / "Documents" / "Graduation" / "Robot"] https://icons8.com/icons

AI Support for Teaching and Learning (at scale)

Matthew D. Beckman
Department of Statistics

February 28, 2026

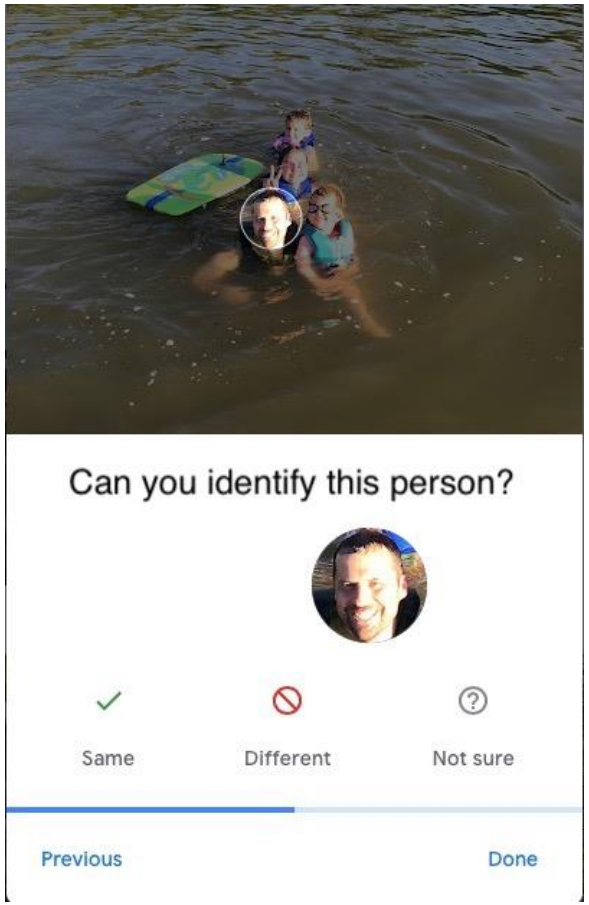
Resource Page

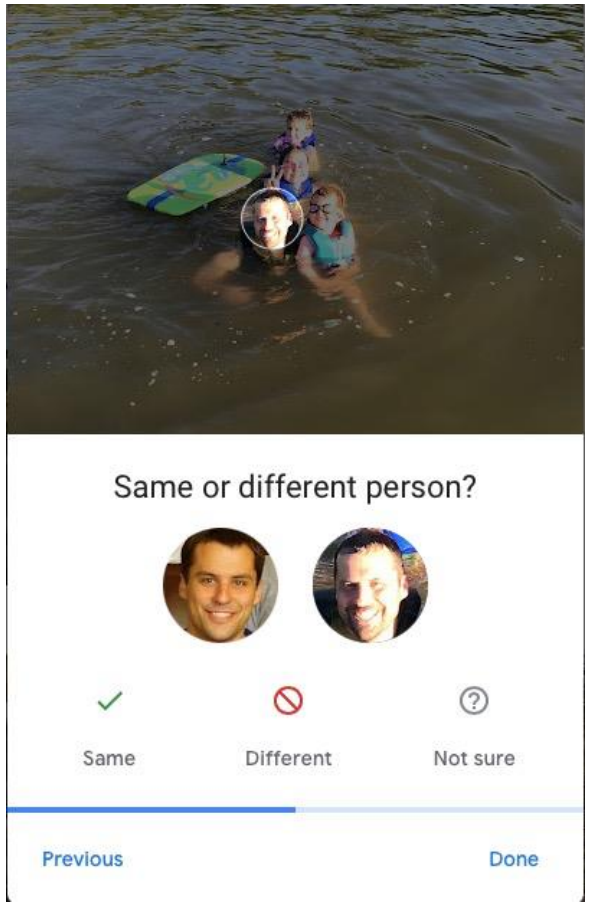
Unexpected Pairings Addressing Big Questions

PennState Eberly College of Science | Ashtekar Frontiers of Science

Backup Slides

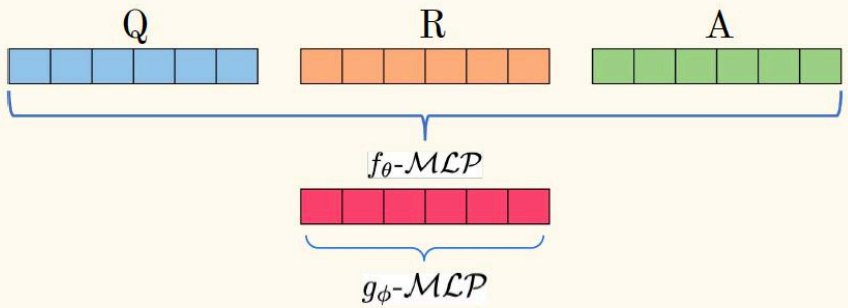
Human-in-the-Loop Intuition





Relation Network Intuition

Motivation for a Relation Network



Q: Susan has samples of 5 different foods. Using only the results of her experiment, how will Susan know which food contains the most sugar? (Gas volume is evaluated by tube)

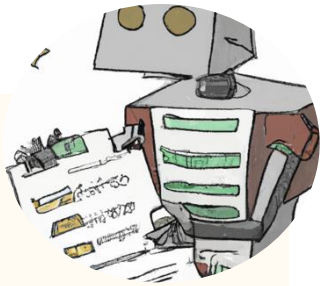
R: Susan should compare the amount of gas in each bag. The bag with the most gas contains the food with the most sugar.

A: Susan will know how much sugar is in the foods by putting each bag in a volume tube. When her finder stops after pushing the top, the bottom of the part she pushes down will be on a number. That number is the milliliters of sugar in the food. Whichever number is the highest, that means that food has the most sugar.

- Many short-answer datasets have triples
  - Question prompt
  - Rubric OR Reference answers
  - Answer from student
- Transformers are less practical
  - Datasets are often relatively small
  - Learning a single vector can efficiently capture relational structure

Credit: Becky Passonneau

Semantic Feature-wise Transformation Relation Network (SFRN)



Answer-State Recurrent Relation Network